Description of TruSight Cancer Panel version 2 (TSCPv2)

TruSight Cancer was developed by The Institute of Cancer Research's Division of Genetics & Epidemiology in collaboration with Illumina® and has been validated for clinical testing in the $TGL_{clinical}$ laboratory.

TruSight Cancer Panel version 2 (TSCPv2) targets 100 genes associated with cancer predisposition, and can be used to screen different subsets of genes depending on the referral type, for example *BRCA1* and *BRCA2*.

The methodology for selection of the TSCPv2 content is attached as an Appendix to this document and the list of genes included on the panel is downloadable from the Mainstreaming Cancer Genetics website. (http://mcgprogramme.com/mcg-info/).

Sample preparation

Each sample is processed separately during the library preparation stage, which involves fragmenting genomic DNA and adding Illumina® adapters and two index tags. The adapters include the sequences required to enable Polymerase Chain Reaction (PCR) amplification, cluster generation and sequencing of each DNA fragment.

Samples are pooled together in groups of 48, then hybridised to the TSCPv2 probes. These are biotinylated DNA probes that specifically target the genomic regions of interest. The biotinylated hybrids are then physically separated using streptavidin beads, and the non-target DNA is washed away. These libraries are then quantified and diluted prior to sequencing.

Sequencing

Sequencing of the prepared libraries is performed on the Illumina® HiSeq2500 in Rapid mode, which uses Illumina® sequencing by synthesis technology. Rapid sequencing runs are carried out using 101 base paired-end reads, with dual indexing. The combination of two index tags allows de-multiplexing of the sequencing reads to a single sample at the data analysis stage. The HiSeq2500 is controlled by the HiSeq Control Software, and during the run quality metrics are monitored using Sequencing Analysis Viewer.

Analysis

TSCPv2 sequence data analysis is conducted through our bespoke pipeline of software tools that:

- convert the raw sequencing images to de-multiplexed .fastq files
- align the reads to the human genome sequence
- call and annotate small variations and large deletions and duplications
- assign quality metrics for each exon
- assign exon failure

Every base in the exon must have at least 50 reads covering the base. An exon which does not fulfil all of these criteria is assigned as a failure. Genes to be analysed for a particular referral type are specified prior to running the analysis, and only data for those genes are outputted for interpretation. The pipeline has undergone extensive evaluation and validation.

Pathogenic mutations and rare variants are additionally confirmed by PCR and Sanger sequencing or Multiplex Ligation-dependent Probe Amplification (MLPA) as appropriate.

All pathogenic mutations and rare variants are reported for every gene tested**.**

Interpretation

Clinical impact is assessed for those pathogenic mutations and rare variants detected by the TSCP pipeline. The TSCP interpretation pipeline has been developed by Clinician-Scientists, combining expertise in genetic cancer predisposition, clinical genetics, molecular genetics, genetic epidemiology, data-mining and large-scale bioinformatics.

The TSCP interpretation pipeline involves expert curation of multiple gene-based and variant-based resources together with expert knowledge of the clinical relevance of variation in the particular gene. Every sequence variation identified is given a clinical classification of one of the following:

- **Pathogenic mutation -** clinical management information for individual and relatives given
- **Variant -** no clinical actions in the individual or predictive testing in relatives recommended
- **Variant requiring bespoke management -** specific recommendation made (e.g. variant is predicted to affect splicing. We require an additional sample for further evaluation).

All reports are authorised by laboratory and clinical personnel.

MCG Mainstreaming Cancer Genetics

Methodology for target selection for TruSight Cancer Panel

The TruSight Cancer Panel was designed to allow analysis of genes and genetic variants associated with predisposition to cancer. It targets germline variants detectable in lymphocyte DNA (rather than the somatic variants that are restricted to tumor tissue).

The panel focuses on genes for which there is already strong evidence of an association with cancer; it does not contain the many genes/variants for which a putative role in cancer has been proposed, but not proven.

The panel focuses on genes that predispose to malignant cancers, although many genes also predispose to non-malignant tumours. A minority of genes are associated predominantly/solely with non-malignant tumours.

We used the following methodology to select the Cancer Panel content:

1) Systematic, expert review of the scientific literature, to identify genes that predispose to cancer. One or more of the following were considered strong evidence of association of a gene with cancer :
   i) Linkage and positional cloning evidence demonstrating segregation of gene mutations in individuals with the cancer phenotype.
   ii) Statistically significant difference in the frequency of the relevant mutation class (e.g. protein truncating mutations or activating missense mutations) between individuals with the cancer phenotype and controls.
   iii) Evidence of *de novo* mutations in individuals with the cancer phenotype.

2) For certain cancer-associated conditions all causative genes were included on the panel, even if some have not been formally associated with cancer. This is because one often does gene testing before an individual has developed cancer, and it is therefore important to be able to analyse all genes associated with the condition.

3) Common variants, typically identified through genome-wide association studies were included if association at $P \leq 5 \times 10^{-8}$ was reported in the Catalog of Published Genome-Wide Association Studies on 06/08/2012 (Hindorff et al).

MCG Mainstreaming Cancer Genetics

We also included 24 polymorphic variants to facilitate sample identification as reported by Fisher et al, comprising 23 common SNPs and a Y-chromosome insertion-deletion (to inform on sample gender).

Our aim is for the panel to be fully comprehensive of all genes/genetic variants that predispose to cancer/cancer-associated conditions. This will require ongoing iteration. Content will be updated as new genes emerge and/or evaluation of existing genes allows their inclusion.

We welcome comments and suggestions about the content, which should be emailed to: mcg@icr.ac.uk

References

Fisher S et al. A scalable fully automated process for construction of sequence-ready human exome targeted capture libraries. *Genome Biol* 12(1), R1 (2011).PMID21205303

Hindorff LA, Junkins HA, Hall PN, Mehta JP and Manolio TA. A Catalog of Published Genome Wide Association Studies. www.genome.gov/gwastudies